

## **Dotplot**

- Open Source; Eclipse Rich Client Platform
- gehostet auf Sourceforge
- Visuelle Darstellung der Ähnlichkeiten von Dokumenten
- Vorbild aus der Genetik, um Ähnlichkeiten von Gensequenzen zu ermitteln

## **Parser Plugin**

Untersucht Struktur von Quelltexten auf Gleichheit => Plagiaterkennung möglich

Besteht aus 4 Hauptkomponenten

### **Converter**

- Haupteinstiegspunkt, wird von Dotplot Framework angestoßen
- Steuert die Erstellung der .mts Datei
- Benutzt dazu den Mapper
- Führt eine interne Typkonvertierung durch

### **Mapper**

- Benutzt den von JavaCC generierten Parser, um Quelltext in Baum zu wandeln
- Dieser Baum wird mit Hilfe des ConfigMappers in eine XML Datei transformiert
- Gibt Referenz auf das XML Dokument im Speicher, an den Converter zurück

### **ConfigMapper**

- Hilfsklasse um vorgegebene Knotennamen des Baumes zu eigenen Namen zuzuordnen
- wird von der Mapper Klasse benutzt
- Konfigurationsdatei leicht erweiterbar, da XML Struktur

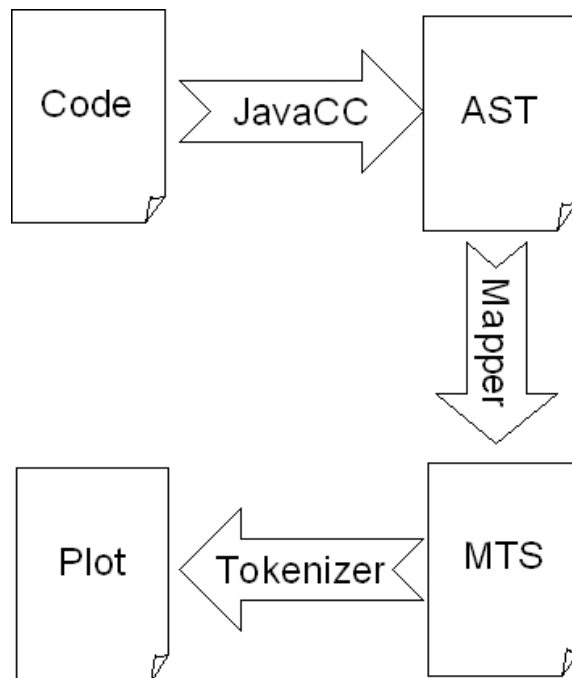
### **Tokenizer**

- Bekommt vom Converter eine Referenz auf XML Datei
- Parsed XML Datei und erzeugt Tokens für den Plotter

## **Diskussion**

- Wie gut ist die Plagiaterkennung mit diesem System?
- Macht das Mappen der Knotennamen überhaupt Sinn?
- Sollten Elemente, mit nur einem Kind in der .mts Datei, gefiltert werden?
- Wie flexibel ist das Parser Plugin in Hinsicht auf eine Metasprache?
- Ist eine gleiche Struktur der Grammatiken für eine Gruppe von Sprachen überhaupt möglich?

## Ablauf



## Klassendiagramm

